Aizen: Automated Big Data Processing, Management and Collaboration

Alon Arad[1], Tim Harris[1], Simon Harris[1], Lori Hathon[2], benjamin Birnbaum[1] and Michael T. Myers[2]

[1]Automated Analytics, United States, [2]university of Houston, United States

Storing, processing, analyzing, and collaborating with large image data sets is a problem in many industries. To address this need we have developed Aizen, a cloud based, big data processing, management and collaboration platform. Users access Aizen via any web browser without requiring any special hardware. Loading large image data sets to Aizen is done by simply dragging the image into the Aizen inbox. Aizen has 2.5D viewing capability for images of any size, with support for TIFF, DICOM, and CZI formats. 2.5D means the view is 2D but all "slices" of a 3D image can be viewed with simple navigation buttons. Overlays with opacity controls are supported. The only parameters that are required for uploading images for analysis data are those that are guaranteed to be known a priori which are used for cataloguing the image(s) so that simple searches of the database can be performed. These would include image modality, instrument ID, and other metadata (sample type, sample location, etc.).

Aizen is built on a fast, robust, reliable (any level of redundancy and failsafe is available) and scalable storage system, so data are an easily searchable asset. Aizen is also built for collaboration. Users can share data sets simply by sending a link. Any user that opens the link will be taken directly to the location the originator was viewing. Comments can be left at any location within images, and regions of interest can be defined, and downloaded at high resolution for analysis in other software tools.

 Aizen is equipped with state-of-the-art machine learning algorithms that are parallelized for the CPU and GPU. This allows the processing of data sets that were previously too large (10's to 100's of GB's) for existing software to even load, much less analyze, to be processed in real time. Aizen reduces runtimes by multiple orders of magnitude on these very large data sets; particularly with multi-GPU enabled algorithms. The adaptive nature of Aizen's machine learning algorithms means that no user inputs are required to achieve optimal results. A full pipeline of entirely automated processing can be performed with nothing more than mouse clicks. Each stage of the processing pipeline is stored and is subsequently available for viewing and further processing or re-processing with manual intervention if required. There is full end to end version control for all actions performed. This allows the entire processing pipeline to be recreated exactly from only the original data.

Recognizing that users will need to use Aizen generated data in external applications, Aizen can output data in TIFF, csv and OBJ mesh formats. This allows, for example, the results of a segmentation to be exported as a mesh, which can then be imported into other software to perform fluid flow simulations. Alternatively, the results of statistical computations can be exported as a csv for use in defining model parameters.

Development of new applications for Aizen is ongoing. In the near term, the results of image analysis will be available to train Deep Neural Networks. This will require the user to simply search and select/define the data sets to be used as training data, run one of the pre-configured DNNs on the training data set and, once completed, to use that DNN model as another built-in processing algorithm. Additionally, upcoming versions will include automated alerting based on defined rules for specified outcomes of processing i.e. simple completion, and potential matches with objects of interest, or calculation of certain statistics that meet criteria such as minimum content of a component of interest.

CrossMark

References

Weninger WJ, Geyer SH, Mohun TJ, Rasskin-Gutman D, Matsui T, Ribeiro I, et al. High-resolution episcopic microscopy: a rapid technique for high detailed 3D analysis of gene activity in the context of tissue architecture and morphology. Anat Embryol. 2006;211(3):213–21.

Arad, A., 2016, **Using image analysis to estimate unconventional rock properties**: AAPG Search and Discovery Article #90255.

Arad, A., M. Madadi, A. G. Christy, A. P. Sheppard, H. Averdunk, and M. A. Knackstedt, 2010,Understanding elastic properties and acoustic anisotropy at the pore/grain scale: Presented at the 51st Annual Logging Symposium, SPWLA.

High-Definition Heart Visualization using Micro-CT Scanning on Experimental Rats
KC Chen, A Arad, ZM Song, DG Croaker - J Clin Exp Cardiolog, 2018

High-definition neural visualization of rodent brain using micro-CT scanning and non-local-means processing
KC Chen, A Arad, ZM Song, D Croaker - BMC medical imaging, 2018